Unsupervised Detection of Contextual Anomaly in Remotely Sensed Data

Qi Liu[a,*], Rudy Klucik[a], Chao Chen[b], Glenn Grant[c,d], David Gallaher[d,e], Qin Lv[b], Li Shang[a]

[a]*Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, 80309, United States*
[b]*Department of Computer Science, University of Colorado, Boulder, CO, 80309, United States*
[c]*Department of Geography, University of Colorado, Boulder, CO, 80309, United States*
[d]*Cooperative Institute for Research in Environmental Sciences, Boulder, CO, 80309, United States*
[e]*National Snow and Ice Data Center, Boulder, CO, 80309, United States*
Email addresses: {qi.liu, klucik, chao.chen, glenn.grant, qin.lv, li.shang}@colorado.edu, david.gallaher@nsidc.org

Abstract

Massive amounts of remotely sensed data are being generated at an unprecedented rate, offering new opportunities for data-driven scientific discovery in the Earth sciences and related domains. However, due to the sheer volume of remotely sensed data and the lack of effective data analytics tools, most data remain in the dark, with little to no quality assurance and limited access to high-level analytical tools. Anomaly detection, which aims to find scenarios that differ from the norm, is of particular importance when analyzing remotely sensed data. However, most previous work has focused on identifying individual anomalies, and required prior knowledge of the ground truth for supervised learning. In this work, we propose an unsupervised anomaly detection framework that requires no prior knowledge and is capable of detecting anomalous events, which we define as groups of outlier objects differing contextually from their spatial and temporal neighbors. Such contextual anomalies can be useful in discovering both hidden quality issues in the data and real natural events of significance. We demonstrate the effectiveness of our framework via Web-based tools developed for visualizing and analyzing such contextual anomalies, using two types of data. The techniques and tools developed in this project are generally usable for a diverse set of satellite products and will be made publicly available with the support of the National Snow

*Corresponding author.

and Ice Data Center (NSIDC).

Keywords: `unsupervised anomaly detection`, `contextual anomaly`, `data quality`, `rare events detection`

---

1. Introduction

Recent advances in remote sensing technology have revolutionized the way remotely sensed (RS) data is acquired, managed, and analyzed (Ma et al., 2015; Rathore et al., 2015). More than 200 on-orbit satellites are currently capturing continuous Earth observations (Ma et al., 2015), offering great opportunities for advancing the scientific understanding of the Earth's systems. However, as the proliferation of these products increases, so does the complexity needed for processing them. The variety of data can vary greatly, even within individual data sets (Li et al., 2016). Therefore, human expert-driven data analysis, a laborious and time-consuming process, remains the mainstream approach for data quality assessment (Isaac and Lynnes, 2003; Gonzalez and Datcu, 2011; Borg et al., 2011) and scientific knowledge discovery (Steffen et al., 2004; Ferguson and Villarini, 2012). The sheer volume and complexity of RS data have hampered adequate quality assessment or higher-level analysis such as anomaly detection. While Earth scientists are very interested in studying anomalies such as climate extremes (Coumou and Rahmstorf, 2012; McCright et al., 2014; Easterling et al., 2000; Muster et al., 2015), finding all such anomalies from massive data sets is challenging. Furthermore, RS data is often contaminated with noise or errors which need to be identified and then either corrected or eliminated. Thus, a high demand exists for effective and generic anomaly detection tools which require minimal involvement of domain experts while having the ability to adapt to diverse data sets. Anomaly detection in RS data is challenging for several reasons. (1) Prior models may not exist for determining what constitutes anomalous data. Additionally, unknown types of anomalies may exist in the data. (2) Remotely sensed imagery is often contaminated with noisy pixels or missing data. (3) The dynamic nature of spatial and temporal variations in multiple frequency channels need to be considered. (4) Due to the high volume and variety of RS data, validated ground truth data sets are not normally available for supervised learning. Additionally, there will always exist unusual anomalies in the data that exceed the expectations or prior knowledge of Earth scientists. Unsupervised approaches are thus preferred. In this work, we propose a clustering-based framework for anomaly detection, which requires no domain knowledge of the data set and enables automated anomaly detection on diverse data sets.

While most previous research has focused on detecting *point anomalies* (Chandola et al., 2009; Gupta et al., 2014; Bhaduri et al., 2011), which are individual data points that are considered globally anomalous (e.g., extreme low temperature or high wind), our work focuses on the less-studied *contextual anomalies* (Sun and Chawla, 2004; Alvera-Azcárate et al., 2012), especially in the dynamics of spatial and temporal domains. Contextual anomalies are relative anomalies under specific contexts. For example, a high air temperature trend in the summer may be normal, but if the same temperature trend occurs during a winter period it could potentially be due to data defects or anomalous atmospheric processes (Matthes et al., 2015; López-Moreno et al., 2014; Bokhorst et al., 2012). Such contextual anomalies are of particular importance in Earth sciences research. And an effective solution for detecting contextual anomalies should leverage both spatial and temporal coherence in localized regions. The assumption is that in a natural environment, pixels in close proximity share similar morphology and evolve gradually over time, while anomalous pixels would have low coherence with their neighbors in space and time.

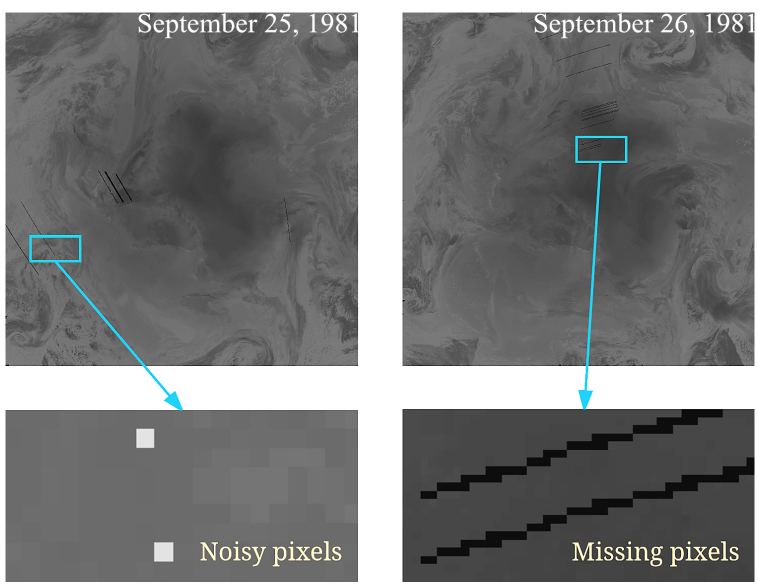

Fig. 1. AVHRR skin temperature data with noise and missing pixels. Examples shown for September 25 and 26, 1981.

One issue to note is the noise or errors in the data. Fig. 1 shows two snapshots of the Advanced Very High Resolution Radiometer (AVHRR) skin temperature data for the South Pole (Chuck et al., 2000, updated 2007). The data values fluctuate from one location to another, as well as over time at the same location. Despite this spatial-temporal dynamic, the data record is also contaminated by random noise from clouds, instrumentation, and missing data. To reduce the bias or disturbance from noisy data when searching for interesting anomalies, we have developed a noisy pixel filtering algorithm and integrated it with the anomaly detection framework.

Besides discovering individual objects ($n \times n$ pixels) that are contextual outliers relative to their spatial-temporal neighbors, it is also helpful to study these outliers collectively as *anomalous events*, which can potentially reveal unusual processes that lead to those outliers in the first place. Such underlying processes can either be systematic errors (e.g., sensor calibration error), which require intervention for quality control, or natural events (e.g., extreme weather condition), which may lead to new knowledge (Xiong et al., 2011; Song et al., 2007). With the knowledge that anomalous behaviors caused by systematic errors or rare natural events can spread to a wide range of regions and last for a long period of time, we aggregate spatial-temporal outliers into anomalous events within a global spatial-temporal context and report those events with a ranking of their importance. Combining all the points above, we have developed a novel clustering-based framework for unsupervised detection of contextual anomalies in remotely sensed data. Our main contributions are summarized as follows.

- The design of an unsupervised anomaly detection framework that (1) requires no prior knowledge of the data set, (2) identifies contextual outliers that differ from their spatial-temporal neighbors; and (3) groups contextual outliers into anomalous events to reveal possible underlying processes.
- Demonstration of the framework's effectiveness via Web-based tools we have developed, using two different types of remote sensing data: SSM/I passive microwave and skin (surface) temperatures derived from AVHRR data.
- Identification and validation of new data quality issues due to systematic or random

errors as well as significant natural events.

This manuscript is organized as follows. Section 2 presents the problem formulation and key notations. Section 3 describes the anomaly detection framework as well as its usage scenarios. Section 4 presents the anomaly detection framework in detail. Section 5 reports our evaluation of the proposed framework and presents case study results. Section 6 gives an overview of related work. Finally, Section 7 concludes this work.

2. Problem Formulation

Climate extremes such as unusual warm and cold events are increasingly attracting the attention of Earth scientists (Matthes et al., 2015; López-Moreno et al., 2014; Bokhorst et al., 2012). In this section, we first introduce the notion of contextual anomalies that can be caused by such extreme events and then formally define spatial-temporal outliers and anomalous events, which are the main focus of our unsupervised contextual anomaly detection framework.

Fig. 2 and Fig. 3 illustrates three types of contextual anomalies that can be caused by unusual natural events, systematic or random errors.

- **Unusual time series snippets**: The Earth observation data usually has obvious cycles (e.g., diurnal or seasonal cycles). In addition, the period of a cycle and the average magnitude of the data in each cycle are relatively stable. However, there exist snippets in a time series that deviate from the stable pattern. For example, Fig. 2 shows a brightness temperature time series from several adjacent pixels. The duration of high brightness temperatures each summer is relatively stable. However, as noted in the figure, in one snippet the high brightness temperature persisted longer than usual, and in another snippet the data value was significantly higher than that of the previous summers. These unusual time series snippets can be caused by either unusual natural events or errors.

- **Level shifting**: Fig. 2 also shows a scenario when the values of a group of adjacent pixels significantly increase or decrease. This type of temporal discontinuity may ap-

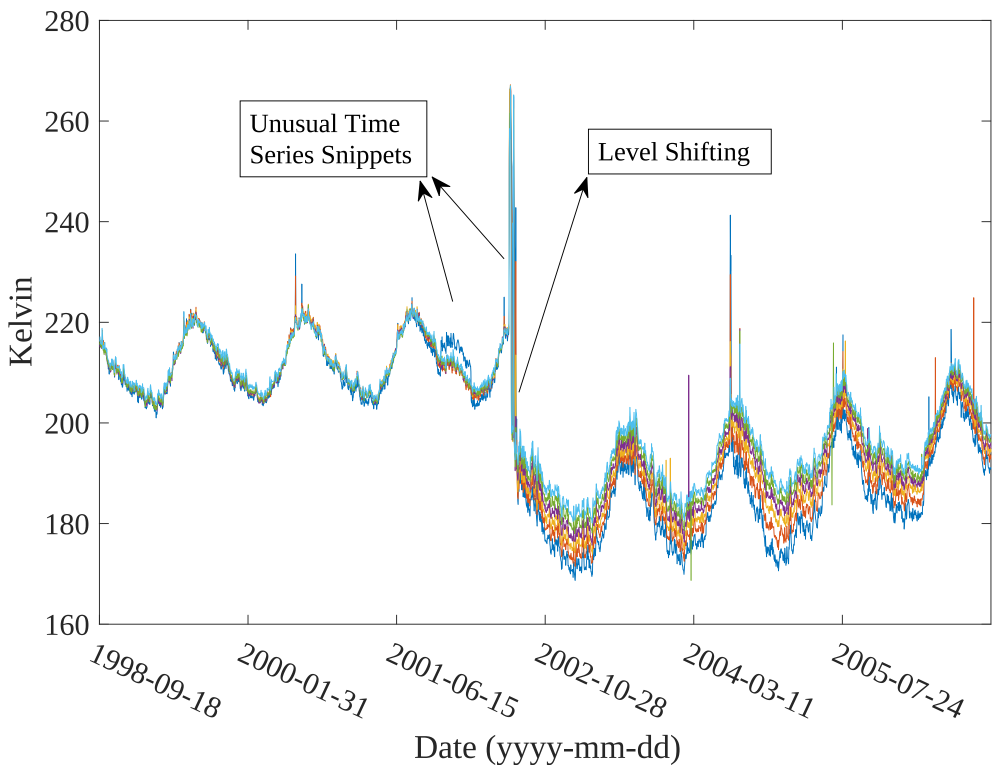

Fig. 2. Examples illustrating unusual time series and level shifting detected in the brightness temperature of several adjacent pixels.

pear normal when viewed spatially at a specific time, and can only be discovered when viewed as a time series at a given location.

- **Local spatial outlier**: A pixel or an object (i.e., a block of $n \times n$ spatial pixels), which appears normal when viewed globally in an image, appears inconsistent when compared with its neighbors. As shown in Fig. 3, pixel A is an outlier with respect to its neighbors, but normal when viewed globally. Pixel B has the same value as pixel A, but is not a local spatial outlier.

These three types of anomalies may all comply with the normal global data range and hence are invisible when using common statistical analyses such as a two standard deviation criteria method against a normal data distribution. Moreover, a level shifting or an unusual time series snippet may be visible only from a temporal perspective. Therefore, in order to detect all these anomalies, both spatial and temporal contextual information is extracted from a pixel's local neighborhood. Specifically, as illustrated in Fig. 4, an object $o^t_{x,y}$ is

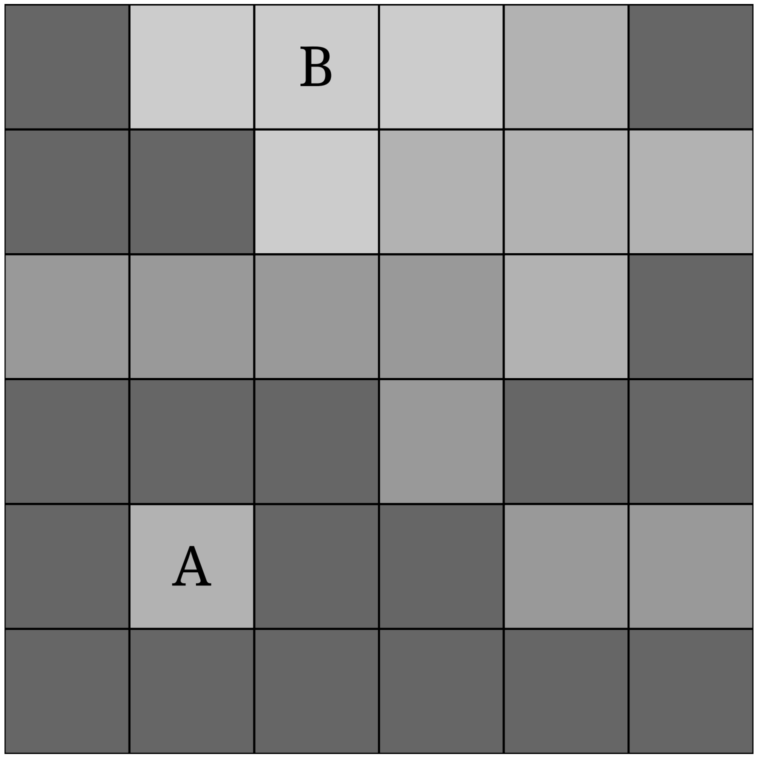

Fig. 3. An example of a local spatial outlier: Pixel A and B have the same value. However, pixel A is considered an outlier because of its behavior with respect to neighboring pixels whereas Pixel B is not because of its coherence with neighboring pixels.

comprised of a block of $n \times n$ pixels at a specific time t, where (x, y) is the location of the top left pixel of the object. Then the spatial neighborhood of the object refers to the eight spatially-adjacent objects at time t, and its temporal neighborhood refers to the set of objects $o_{x,y}^{t'}$ where $t' \in [t - T, t + T], t' \neq t$ and T is the window size parameter. We now define spatial-temporal outliers (ST-Outliers) and anomalous events as follows.

Definition 1. ST-Outliers: Given a set of objects $O = \{o_{x,y}^{t}\}$. An object $o_{x,y}^{t} \in O$ is considered an ST-Outlier if its non-spatial-temporal features differ significantly from its spatial or temporal neighbors in O.

The non-spatial-temporal features represent an object's original physical value such as temperature, or derived values, e.g., temperature difference, temperature correlation and so on. Because ST-Outliers can emerge as a group due to the same natural event or systematic error, we also define anomalous events.

Definition 2. Anomalous Events: Let D be a set of objects that are ST-Outliers, and r is an anomalous event that consists of a group of objects from D. The objects in r, which are spatially and temporally correlated, behave significantly different from the other objects in D in terms of non-spatial-temporal features.

In the following section, we give an overview of our anomaly detection framework along with the web-based tools we developed that enable users to visualize and analyze the detected anomalies.

3. Anomaly Detection Framework

Our anomaly detection framework takes as input a time series of satellite images. Each image is usually processed at either the pixel level or the object level (Hussain et al., 2013). Noise or missing data usually appear as random, discontinuous pixels. However, interesting anomalies that represent natural events or systematic errors may often appear as a collection

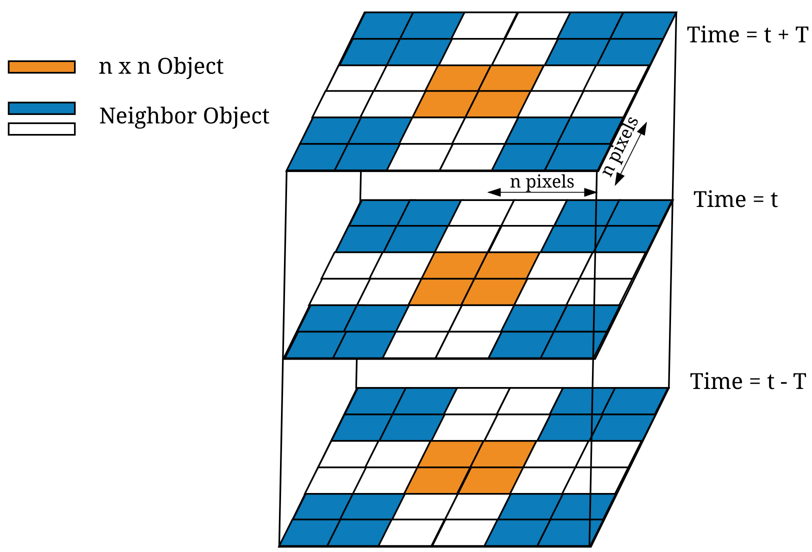

Fig. 4. An illustration of objects and local spatial-temporal neighborhoods. Each object is defined as a block of 2×2 pixels at a specific time. For the orange object at time t, its spatial neighbors include the 8 adjacent objects surrounding at time t (4 in blue and 4 in white), and its temporal neighbors are the other orange objects within the time range of $[t - T, t + T]$.

of adjacent pixels. As such, our proposed anomaly detection framework consists of both pixel-based and object-based analysis.

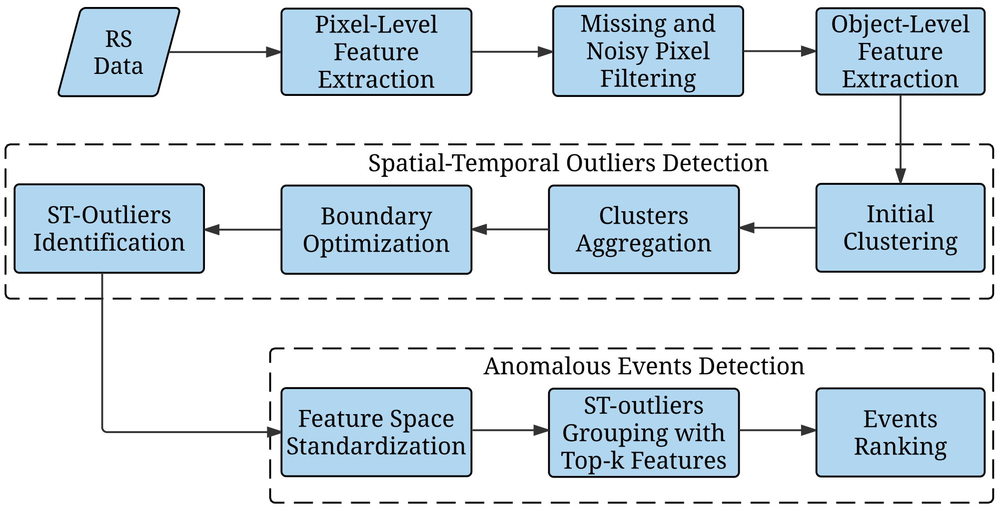

Fig. 5. An overview of the anomaly detection framework.

3.1. Overview

Fig. 5 gives an overview of our proposed anomaly detection framework, which consists of four main steps:

1. **Missing and noisy pixel filtering**. RS imagery data may be contaminated by missing and/or noisy pixels, which would lead to a skewed data distribution. Thus, this step is designed to improve the quality of detected anomalies, which can be used as either an independent tool for data cleansing or integrated into the anomaly detection process, as shown in our framework.
2. **Object-level feature extraction**. Each object consists of one or more pixels ($n \times n$). In order to capture the anomalous behaviors of an object, the radiometric values are extracted for each object, which are then compared with its neighbors in space and time to detect contextual anomalies.
3. **ST-Outliers detection**. An object that has either low spatial or temporal coherence with its neighbors is identified as an outlier. This is accomplished through an unsupervised, clustering-based process.

4. **Anomalous events detection**. In this step, we further group ST-Outliers that share similar anomalous behaviors (i.e., spatially and temporally correlated) as an anomalous event in order to help discover the underlying anomalous process, be it a natural event or systematic error.

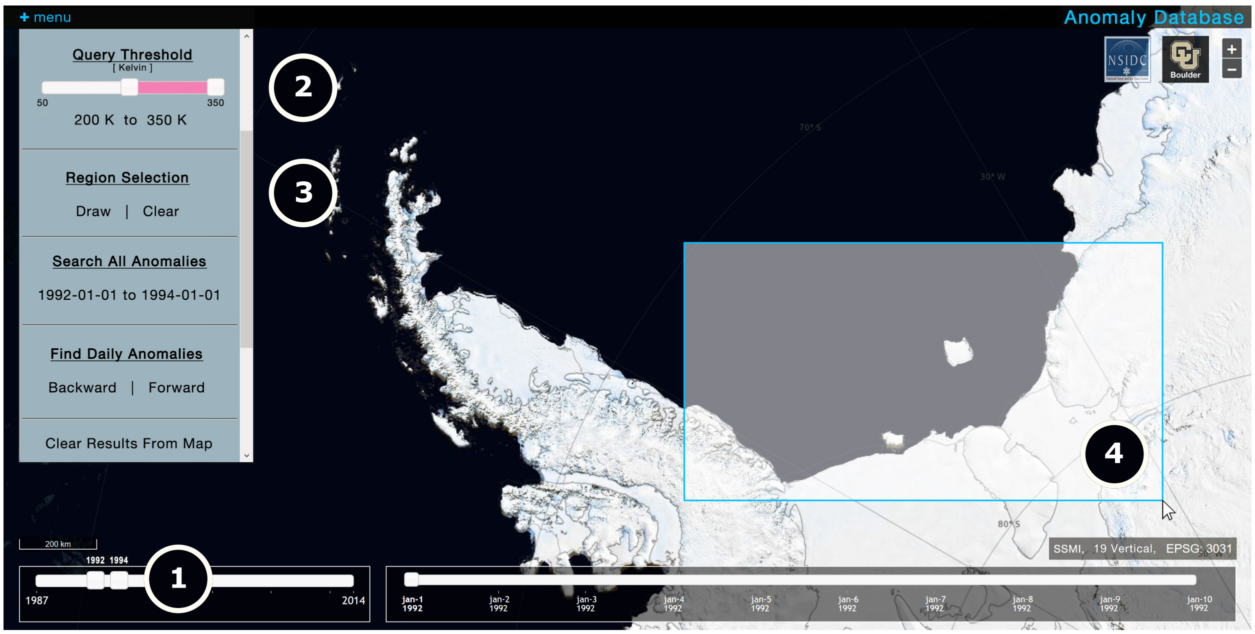

Fig. 6. Web-based user interface: Overall layout and key steps showing how anomalies are located.

3.2. Usage Scenarios

Next, we describe the web-based tools that enable interactive exploration and analysis of anomalies by interfacing with a database that contains all anomalies detected by our framework.

User Interface. Fig. 6 shows the web-based user interface (UI), which features a set of tools and enables data exploration using an intuitive mapping interface. The UI comprises of several key features that allow users to quickly select a set of parameters that include information such as the sensor type (e.g., SSM/I vs. AVHRR), frequency band (e.g., 19 GHz, 22 GHz), and polarization (e.g., vertical or horizontal), as well as select a sub region within the map to search. Users can then explore their results in one of two ways. The first

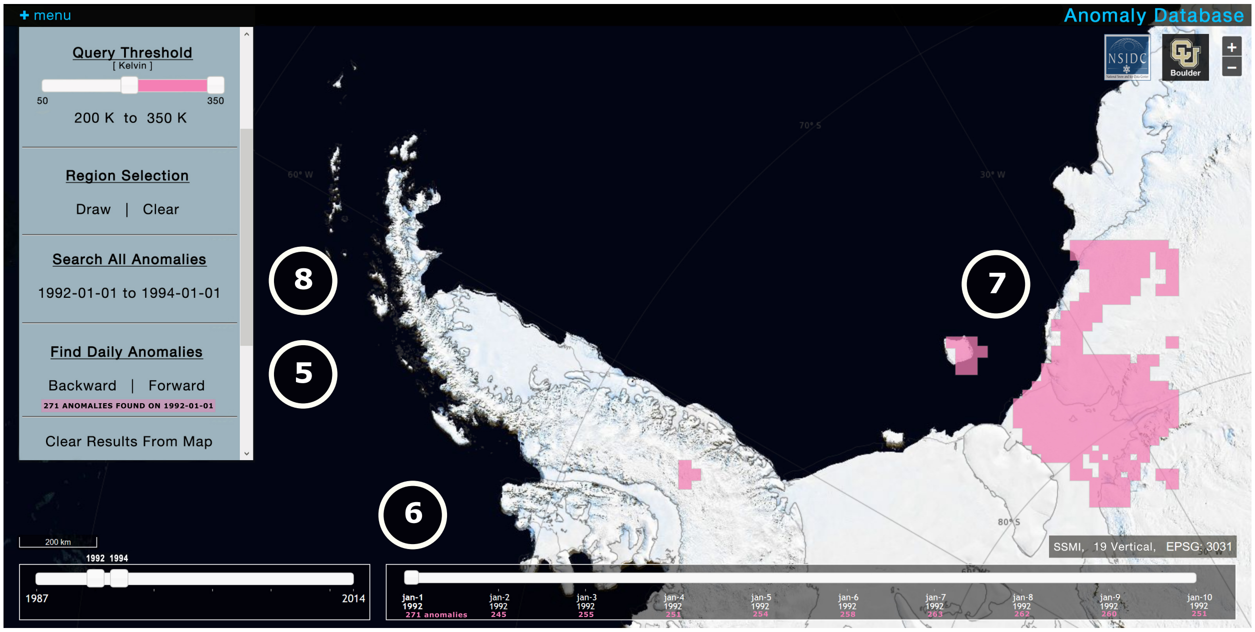

Fig. 7. Usage example: Querying the Weddell Sea and coast for anomalies.

approach allows query results to be visualized day by day, where the daily anomalies and their associated metadata are used to help the user grasp dynamic changes within a particular region. The second approach is by searching for an aggregation of the data within the specified time frame. In this way, results are collected for each pixel and can be displayed to show information such as the average pixel value or the frequency with which anomalies occur at each location.

Usage Example. We examine a coastal region to demonstrate how the tools can be used for exploring anomalies; the steps described are illustrated in Fig. 6 and Fig. 7. First, the user selects a date range of 1992 to 1994 of the South SSMI 19 vertical data set, refining the query to search for anomalies above 200 Kelvin with the slider. Next, a rectangular region is selected from anywhere in the Antarctic region; the user opts to search an area within the Weddell Sea, drawing out a rectangle to partition the specific sub region they are interested in. Finally, the user then selects a query that will parse anomalies so that they can be reviewed along the daily timeline. After the queried results are returned to the interface, they are overlaid within a layer of the map where the user has control over a temporal investigation of the data, allowing them to traverse forwards and backwards along the timeline to review results. Looking at the results, we can see that there are a significant number of events concentrated within the Filchner Ice Shelf extending up toward the Brunt Ice Shelf. This may prompt the user to refine his/her search to look at an aggregation of all anomalies during the specified two-year period, allowing the user to see where those anomalies are concentrated most, and the brightness temperatures associated with each pixel.

4. Anomaly Detection Algorithms

In this section, we describe in detail the key steps in the anomaly detection framework.

4.1. Missing and Noisy Pixel Filtering

As mentioned earlier, most RS imagery contains missing and/or noisy pixels, which need to be properly identified and labeled. This information is then applied in the process of object-level feature extraction in order to reduce the bias introduced by those pixels. Additionally, we record noisy pixels in the anomaly database as random errors for users' reference. Missing pixels are easy to handle since a missing pixel's value is usually set to a special fill value such as 0. Hence, we focus primarily on filtering noisy pixels. Some of the noisy pixels are outside the normal data range (i.e., clear errors), and can be filtered easily using a threshold. However, some of the noisy pixels are within the normal range but obviously "wrong" when compared with their neighbors. To address these scenarios, we have designed a noisy pixel filtering algorithm, which can detect both types of noisy pixels (outside normal data range or not) in two steps: (1) identify objects which contain potential noisy pixels and (2) identify actual noisy pixels in each object.

Detect Potential Noisy Objects. In order to reduce the size of the feature space and improve computation efficiency, we first divide an image into objects of $n \times n$ pixels each. For each object, we extract features such as the absolute maximal difference between every two adjacent pixels. Fig. 8 shows a distribution of the absolute maximal difference between every two adjacent pixels in an object from an AVHRR image. The cutoff value in the distribution is around 38. Empirically, this cutoff point can be utilized to find objects that contain potential noisy pixels. However, a fixed threshold is not generic for all images in the AVHRR data or for other data sets. Therefore, we have developed a clustering-based method to automatically determine the "cutoff" threshold. The feature data is first clustered, then

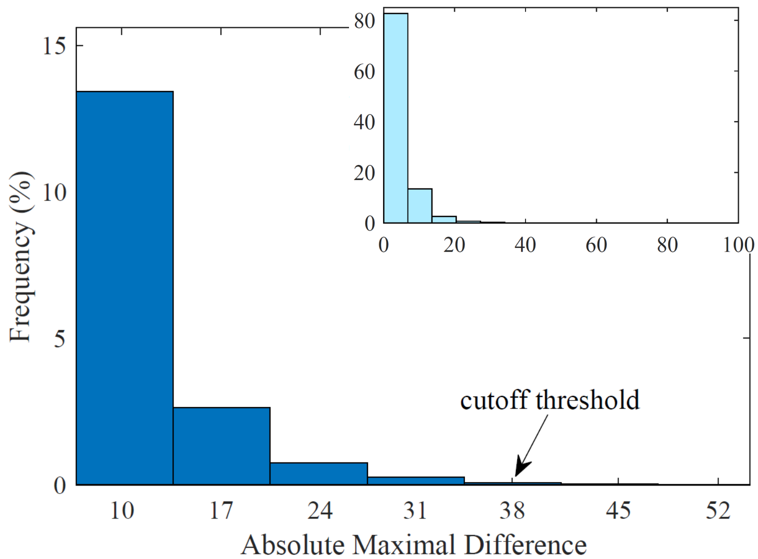

Fig. 8. Histograms showing the frequency distribution of absolute maximal differences between adjacent pixels in each object. The main plot (blue) highlights a subset of differences from the total distribution (cyan). The threshold for determining whether an object contains potential noisy pixels can be visually selected such that the probability of the absolute maximal difference converges toward zero.

the second order difference of cluster centroids (sorted in the ascending order) is computed. As illustrated in Fig. 9, the first peak of the second order difference (shown in red) is detected, and the cluster centroid value (shown in blue) right after the peak is set as the threshold to detect objects that contain potential noisy pixels.

Identify Noisy Pixels. For each object detected above, we divide its pixels into two groups based on the similarity of their features, and pixels in the smaller group are identified as noisy pixels. However, there is one potential issue. As illustrated in Fig. 10, the pixels (P1 and P2 in blue box) from the edge of the Antarctica Peninsula are identified as noisy pixels in object B, but it looks normal in object A. In order to handle pixels that are located around edges or dynamic regions (e.g., ocean), we compute the absolute difference between each noisy pixel candidate and its neighbors. If a pixel is similar to the majority of its neighbors, the pixel is not noisy since we assume noisy pixels are random and do not occur together.

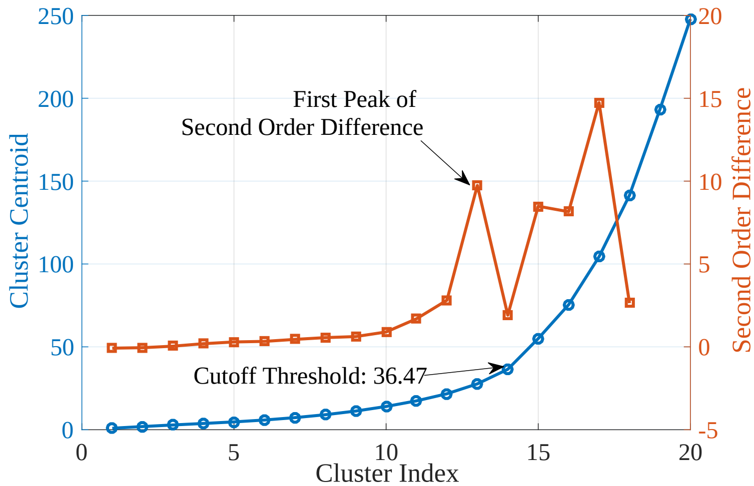

Fig. 9. A method for automatically determining the cutoff threshold. Step 1 finds the first peak of the second order difference (shown in red), here the cluster index is 13. Step 2 checks the cluster centroid series (shown in blue) to find the value 36.47 at index 14, that value is then used to detect objects containing potential noisy pixels.

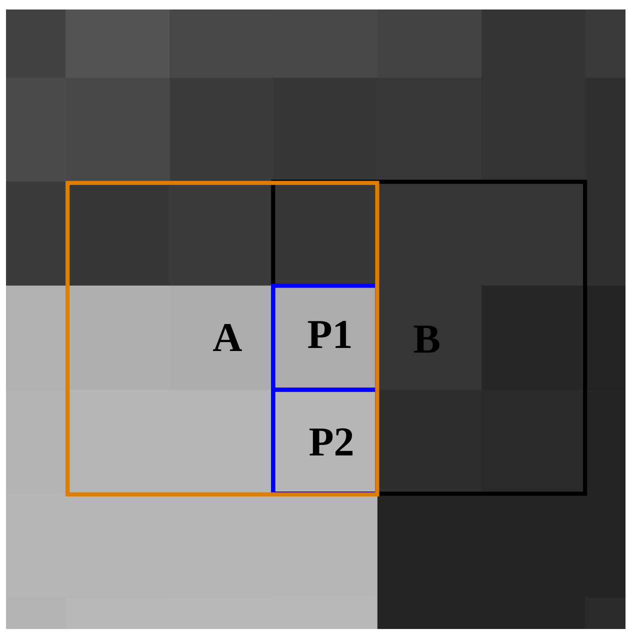

Fig. 10. One potential issue when identifying noisy pixels near edges and dynamic regions. Pixels P1 and P2 are observed as normal within object A (orange 3×3 pixels block), but are identified as outliers in object B (black 3×3 pixels block) due to the sharp transition at the edge of the Antarctic Peninsula.

4.2. Object-Level Feature Extraction

In this step, instead of extracting pixel-level features to emphasize the internal pixel-to-pixel variance of an object, we extract object-level features to describe an object with respect to its neighbors. Specifically, we extract three types of features for each object.

- Basic features: Mean and standard deviation of the object's $n \times n$ pixels.
- Spatial features: Correlation and difference between the object and each of its eight neighboring objects to capture the spatial dynamics of an object against its spatial neighbors.
- Temporal features: Mean and standard deviation difference between the object and each of its $2T$ temporal neighbors (i.e., $[t-T, t+T]$).

By transforming the satellite image time series into a feature space, we achieve three goals. (1) Reduce the impact from noisy and missing pixels: For instance, a missing pixel can be ignored by computing the mean and standard deviation of an object that contain multiple pixels; but if all pixels of an object are missing pixels, the object is abandoned. (2) Leverage the spatial features to avoid the problem of spatial heteroscedasticity, i.e., the local distribution of the data is not uniform at different locations in an image. (3) Help remove the impact from cyclic patterns and highlight local outliers.

4.3. ST-Outliers Detection

Satellite imagery data generally contain some cyclic patterns such as seasonal cycles. It is thus reasonable to assume that the data follows Gaussian mixture models (GMM). However, the number of clusters K requires prior knowledge and is very difficult to determine. Our problem is even more complicated because the exact percentage of outliers is unknown. To address these issues, we propose an extended Expectation-Maximization (EM) algorithm. Fig. 12 illustrates the essential steps of our algorithm for detecting spatial-temporal outliers: cluster initialization, cluster aggregation, and boundary optimization. The rationale behind this approach is to focus on capturing the normal patterns, and treat small clusters and stand-alone data points as outliers, because normal patterns occur more frequently than outliers and belong to denser clusters. Next, we describe the three steps in more detail.

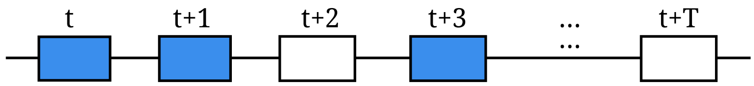

Fig. 11. A time series of objects at location (x, y) spanning time t to $t+T$. Normal objects are shown in blue and outlier objects are shown in white.

Cluster initialization. In this stage, we detect outliers from all objects $\{o^{t_0}_{x,y}, \ldots, o^{t_T}_{x,y}\}$ at location (x, y) within the maximal temporal span, as shown in Fig. 11. Note that spatial-temporal outliers can be captured because both spatial and temporal features are also used here. For example, if one object is significantly different from its spatial neighbors, this object should be significantly far away from the others in the spatial feature space so that it will be assigned to a small cluster or be a stand-alone object. Moreover, since the exact number of clusters is unknown, we choose a relatively large cluster number ($K = 5$ as shown in Fig. 12 (a)) to perform the multivariate EM algorithm (Andersson, 1958; Lee and Scott, 2012) in the object feature space. However, if K is too large, similar objects may be assigned to different clusters.

Cluster Aggregation. To address this issue, we merge clusters with very similar statistical models. For example, in Fig. 12, the top-left two clusters shown in red and green in step (a) are merged into the larger red cluster in step (b). Let $\{\mathcal{C}_1, \ldots, \mathcal{C}_K\}$ be the initial clusters,

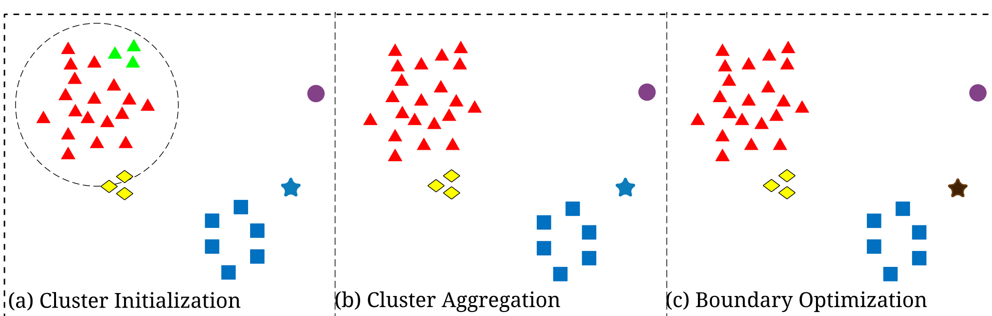

Fig. 12. Illustration of the three key steps of the proposed ST-Outliers detection algorithm. Object shapes represent the ground truth of different types of objects, and object colors indicate the clusters they belong to. In this example, step (a) identifies five different clusters; step (b) merges the green cluster into the red cluster; and step (c) separates the brown star object from the blue cluster.

and $\{\mathcal{N}_p(\mu_1, \Sigma_1), \ldots, \mathcal{N}_p(\mu_K, \Sigma_K)\}$ be the corresponding statistical models, where p is the dimension of the feature space, and μ, Σ are the mean vector and variance matrix respectively. We define the impact domain of model $\mathcal{N}_p(\mu_i, \Sigma_i)$ as $A^i_\alpha = \{x \mid \Pr[(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i) \leq \epsilon] \leq 1-\alpha\}$. Intuitively, a greater α means a tighter impact domain. Based on the relation between $\mathcal{C}_i$ and A^j_α for any (i,j), we have the following three situations:

$$\mathcal{C}_i \subseteq A^j_\alpha \tag{1}$$

$$\mathcal{C}_i - \mathcal{C}_i \cap A^j_\alpha \neq \phi \tag{2}$$

$$\mathcal{C}_i \cap A^j_\alpha = \phi \tag{3}$$

The three situations are presented in Fig. 12 (a). Assuming the red triangles cluster represents $\mathcal{C}_j$, its impact domain A^j_α is within the dashed circle. In the case of 1, assuming $\mathcal{C}_i$ is the green triangles cluster, then $\mathcal{C}_i$ and $\mathcal{C}_j$ can be merged together, because with high probability, the observations in $\mathcal{C}_i$ may be from statistical model $\mathcal{N}_p(\mu_j, \Sigma_j)$. However, in the case of Eq. 3, assuming $\mathcal{C}_i$ is the yellow rhombuses cluster, then the probability of these two clusters are from the same statistical model is very low and cannot be merged. To deal with the remaining case of 2, we define a new statistic W:

$$W = \frac{|\mathcal{C}_i - \mathcal{C}_i \cap A^j_\alpha|}{|\mathcal{C}_i|} \tag{4}$$

If W is larger than a given threshold, then we merge the i^{th} and j^{th} clusters. In this way, The situation in (1) becomes a special case of the situation in (2).

Boundary optimization. After cluster aggregation, we reestimate the statistical models based on the updated clustering results, and then optimize the boundary for each cluster. We remove an object from a cluster if the object does not follow the cluster's statistical model. For example, as shown in Fig. 12 (c), the blue hexagon is removed from its original cluster and becomes a stand-alone object. To do so, we test every object in the i^{th} cluster whether it is from population $\mathcal{N}_p(\mu_i, \Sigma_i)$.

Outlier Identification. We repeat the procedure in cluster aggregation and boundary optimization until there is no further change. After that, the small clusters and isolated objects are considered as spatial-temporal outliers. Please note that the underlying assumption is that the percentage of anomalous data in the whole data set is quite low.

4.4. Anomalous Events Detection

Usually, for domain experts, the ultimate goal of anomaly detection is not identifying the individual outliers, but to find out the underlying processes that cause those outliers. Thus, instead of raising an alarm for every single outlier, it is much more valuable to provide an ordered list of anomalous events along with their specific rankings of importance or level of interest. We accomplish this in the following three steps: feature space standardization, ST-Outliers grouping, and events ranking.

Feature Space Standardization. In this step, we use z-score to compute the standardized score for each type of feature in the feature space:

$$F_{stad} = \frac{F - \mu_F}{\sigma_F}, \tag{5}$$

where μ_F and σ_F are the mean and standard deviation of the feature, respectively. This step is needed because each feature type has a different value range. By using a z-score normalization, all features now fall within the same value range, thus allowing us to compare/group outlier objects using the top-k most significant features.

ST-Outliers Grouping and Events Ranking. In this step, we sort the feature vector for each outlier by the absolute value of F_{stad}, then group the outliers as illustrated in Fig. 13, where the outliers in the same group have the same top-k categorical features. For each group, we merge every outlier and its spatial and temporal neighbors into an event until there is no further change. The intuition behind this grouping strategy is that the impact of an underlying anomalous process usually spans a continuous time period, and from the spatial perspective, more than one object is affected. Finally the events are ranked by the total number of outliers in each event and reported through the web-based UI.

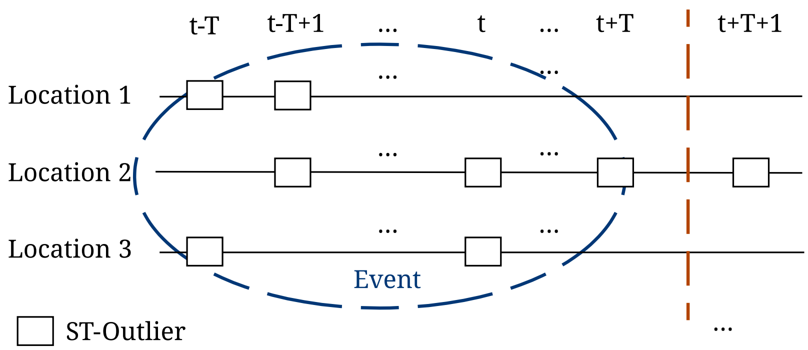

Fig. 13. An example of grouping ST-Outliers into anomalous events. ST-Outliers from three locations are detected at different time points and are similar in their top-k features. The outliers that occur within the same time window $[t-T, t+T]$ are then grouped together as a single event.

5. Results and Discussion

In this section, we first evaluate the performance of the proposed anomaly detection framework with experiments carried out on two data sets: skin temperature derived from AVHRR data and DMSP SSM/I Daily Polar Gridded Brightness Temperatures. The details of parameter settings, and case studies of AVHRR and SSM/I data are discussed. Then, the computational efficiency of the proposed method is analyzed both theoretically and experimentally.

5.1. AVHRR Data

We used the South Pole AVHRR skin temperature data from July 24, 1981 to June 30, 2005. During the process of creating the anomaly database we discovered that the AVHRR data is heavily contaminated with noise; this data set thus served as a good test case for assessing the performance of our noise pixel filtering algorithm.

Parameter Settings. As mentioned in the algorithm design, it is inefficient to utilize clustering with individual pixels due to the large number of pixels in satellite images. Instead, we divide each image into objects of size $n \times n$ pixels to identify and filter out objects which could potentially contain noisy pixels. We experimented with different n values, and setting $n = 3$ achieves a good balance between accuracy and efficiency. The assumption is that within each 3×3 object the variance should be small. Once the absolute maximal difference is computed for each object, the data is transformed into a feature space where the algorithm described in Section 4.1 is applied.

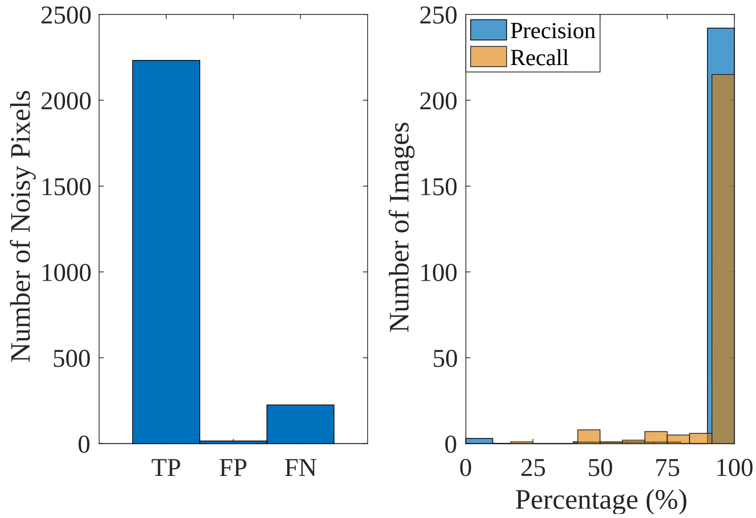

Fig. 14. Noisy pixel filtering in the AVHRR data set. Results show that our algorithm correctly identifies most of the noisy pixels (Left) and achieves high precision and recall for most of the images (Right).

Algorithm Performance. To assess the accuracy of the noise pixel filtering algorithm, we can validate the detected pixels visually using a random sample of 10% of the AVHRR images. We then quantified the performance using two widely-used pattern recognition metrics: *precision* and *recall*.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

True positive (TP) is the number of noisy pixels that are correctly detected as noise. False positive (FP) is the number of normal pixels that are incorrectly identified as noise by the algorithm. And false negative (FN) is the number of noisy pixels that are incorrectly classified as normal. Fig. 14 shows the distribution of each metric evaluated for the data set. The average *precision* and *recall* are 98.1% and 95.7%, respectively. This indicates that our noise filtering algorithm is effective, which can be used for data quality control and filtering, and can help reduce the bias introduced by such noisy pixels in the anomaly detection process.

5.2. SSM/I Data

SSM/I data is a primary resource for estimating sea ice concentrations and classifying sea ice types. While the data set has been continuously collected for nearly 30 years, from July 9, 1987 to June 30, 2015, it has been distributed without a thorough quality assessment. New data defects have been discovered by our framework and confirmed by specialists. Here, we use a case study from the North Pole data set to demonstrate the effectiveness of our approach.

Parameter Settings. For each image in the SSM/I data set, an object is defined as a 2×2 block of pixels, i.e., $n = 2$. Here we use a smaller block size because SSM/I data has a lower resolution than that of AVHRR data (25km vs 5km), and $n = 2$ is the minimum requirement for removing the impact of missing or noisy pixels. With a vector size of four (four pixels in each object), we compute the spatial correlation between an object and its eight spatial neighbors. The temporal neighborhood spans 2 days before and after each object (i.e., $T = 2$) to help smooth out dynamic attributes such as clouds, which usually pass through the area in 1 to 2 days. The temporal neighborhood of 5 days can therefore reduce random noise without filtering out real, dynamic, or periodic fluctuations in the time series. For each object, we extract six features: the mean and standard deviation of the pixels, a spatial correlation vector and a difference vector between the object and its eight neighbors, and the mean and standard deviation between the object and its temporal neighbors (+/- 2 days). The object features are then used for the detection of ST-Outliers and anomaly events.

As the inherent number of models of the data is uncertain, Silhouette coefficients (Rousseeuw, 1987) are used to evaluate the clustering's performance for differing initial conditions. Because the clustering quality is positively related to the Silhouette coefficient, the number of initial clusters is chosen where the Silhouette coefficient reaches a maxima. Fig. 15 shows how the Silhouette coefficient changes with varying number of initial clusters with random samples of 10% of the SSM/I data. In addition, we used 10% as an upper bound for the total number of outliers. As with a relatively larger boundary, we maintain a high potential for

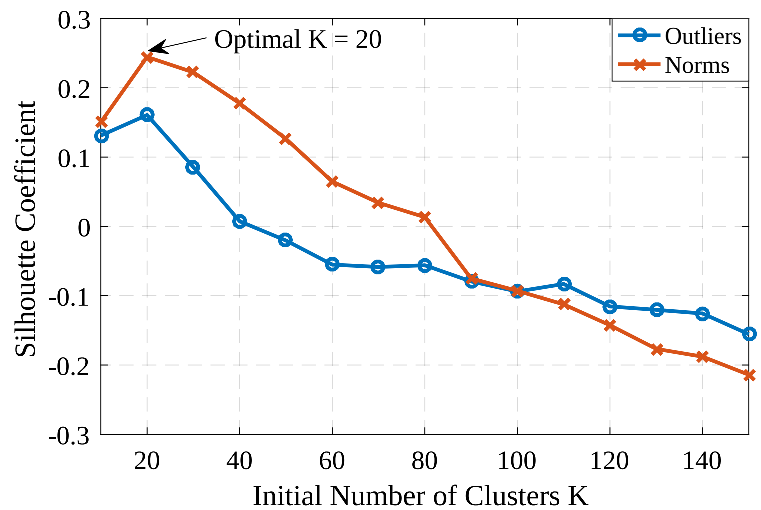

Fig. 15. Choosing the initial number of clusters K. Silhouette coefficients are computed for the normal and outlier clusters using varying numbers of initial clusters. The optimal number of clusters $K = 20$ is selected when Silhouette coefficients reach a maxima for both normal and outlier clusters.

including all anomalous events in the database. We aggregate outliers from the smallest size of cluster until the total number of outliers exceeds 10%. Thus, the resulted total outliers can be equal or less than 10%, which depends on the real partitions of norms and anomalies.

Table 1: List of Top Ranked Anomalous Events

Event Duration	Category
1990.01.01-1991.12.31	Sensor Failure
2012.07.09-2012.07.12	Natural Event
2012.07.27-2012.07.29	Natural Event
2002.06.27-2002.06.29	Natural Event
2003.10.24-2003.10.25	Natural Event
2011.02.01-2011.02.04	Unknown
2010.09.02-2010.09.04	Systematic Error

Case Studies. Table 1 shows a partial list of top-ranked anomalous events discovered and reported by our framework. Because no ground truth exists for this data, we collaborated with other geoscientists and studied previous literature for the region to identify several of the most significant anomalous events; these events exhibited systematic error or evidence of natural events to help validate our technique. Here we discuss several of those events in detail.

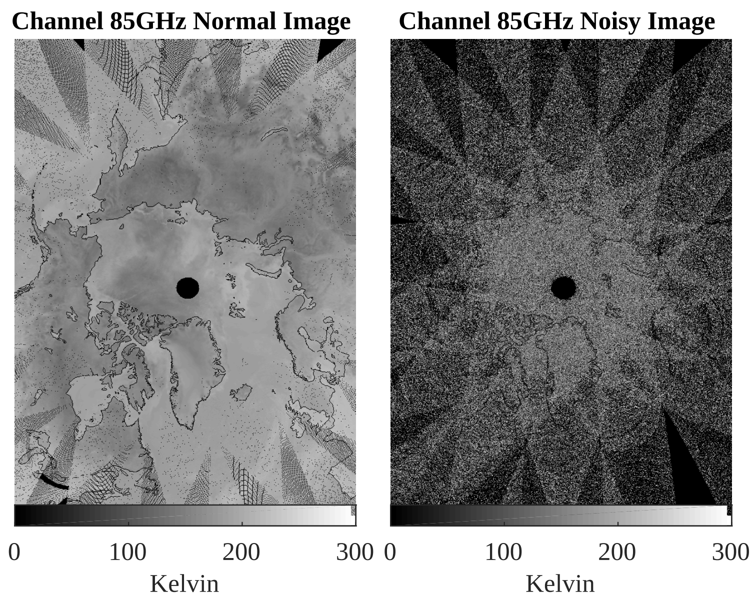

Fig. 16. SSM/I Data Defect: random noise within the image due to sensor failure.

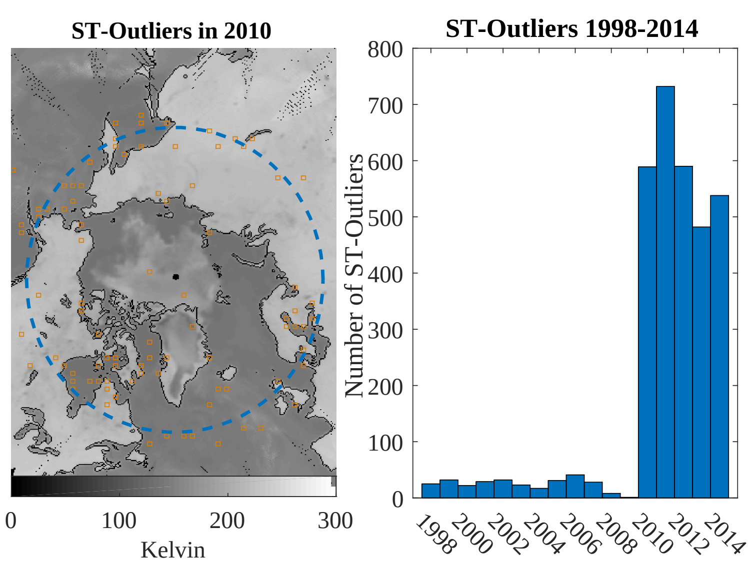

Fig. 17. A systematic error from 2010. (Left) The majority of ST-Outliers were detected around coastal regions. (Right) A significant shift in the number of ST-Outliers beginning in 2010.

Event 1: The first event was found within the 85.5 GHz channel. The 85.5 GHz vertical channel exhibited a degradation in signal from January 1, 1990 to December 31, 1991, while the horizontal channel degraded between January 1, 1991 to December 31, 1991. The origin of the event was a sensor failure. All the images collected during that period were corrupted with random noise, as shown in Fig. 16. This data defect could significantly affect prior analysis and computation of the region's climatology. While only part of the defect (degradation in 1991) was documented (Maslanik and Stroeve., 2004, updated 2016), our algorithm was able to uncover new errors within the 85.5 GHz channel. The issues with the 1990 data were reported to the NSIDC to help alert users.

Event 2: A sharp increase in the frequency of anomalies was discovered following 2010. The left image in Fig. 17 shows the spatial locations of ST-Outliers (orange squares), mainly detected by temporal mean and standard deviation features in 2010. As seen in the figure, the

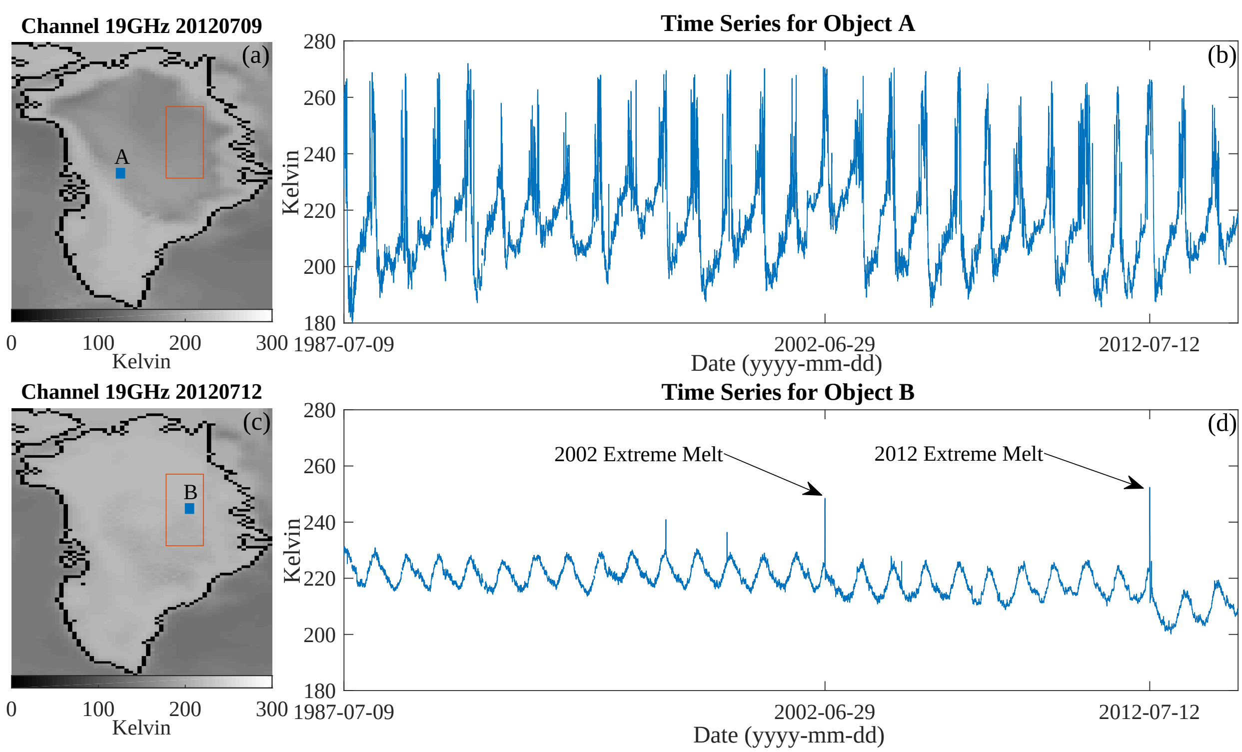

Fig. 18. SSM/I anomalous event: Summer extreme melt in 2002 and 2012. The red boxes in (a) and (c) represent regions which rarely melt. (b) The time series for object A, melting occurs regularly every summer. (d) The time series for object B, which was impacted by the extreme melt events in both 2002 and 2012.

majority of outliers are located around coastal areas. We determined that this surge of events was due to an inconsistency between measurements from the sensors on DMSP satellites F13 and F17 (where data from the F13 sensor was used until 2010, then transitioning to F17). The NSIDC conducted an inter-comparison of the F13 and F17 data where products from the two sensors overlapped. Similar to our findings, larger differences, sometimes up to 10 K, were found in regions of sharp gradients of brightness temperature, usually around coastlines and sea ice extents (Maslanik and Stroeve., 2004, updated 2016). In addition to the discovery of this systematic error, we were also able to generate a detailed report on the spatial-temporal locations of each outlier for the event. This last product could potentially accelerate the quality control process.

Event 3 and 4: Events from 2002 and 2012 are top ranked, consistent with two extreme melt events that occurred during those years (Nghiem et al., 2012; Steffen et al., 2004). As shown in Fig. 18 (a) and (c), the region outside of the red box regularly melts during the summer months, while in 2002 and 2012, that melting process abnormally expanded into part of the region within the red box. Our algorithm effectively detected the locations and dates of regions that normally would not melt when they exhibited abnormal behavior in 2002 and 2012. Fig. 18 (d) shows an example time series from one of those locations (object B) between 1987 and 2015. The brightness temperature reveals a sharp increase during the summers of 2002 and 2012. In Fig. 18 (c), most rare melt objects are located in the red box. In addition, all rare melt locations detected were found to have averaged 5 melt days over the years. Thus, our framework can provide a way to explore potential interesting rare events without manually sifting through the data.

Event 5: Besides the widely known extreme melt events in 2002 and 2012, the algorithm also detected an unusually warm event during October 2003. This event caused one location, which typically would begin to freeze during this time, to experience about an extra month worth of melt days. Fig. 19 shows this unusually long melt event in October 2003. From the time series in this figure, spanning 2002 to 2004, the brightness temperatures of these three adjacent regions normally would reveal a mean value which decreases during the month of October. In 2003 though, there were sharp changes, reaching a maxima one would expect during the summer. Our algorithm accurately captured this event with the presence of

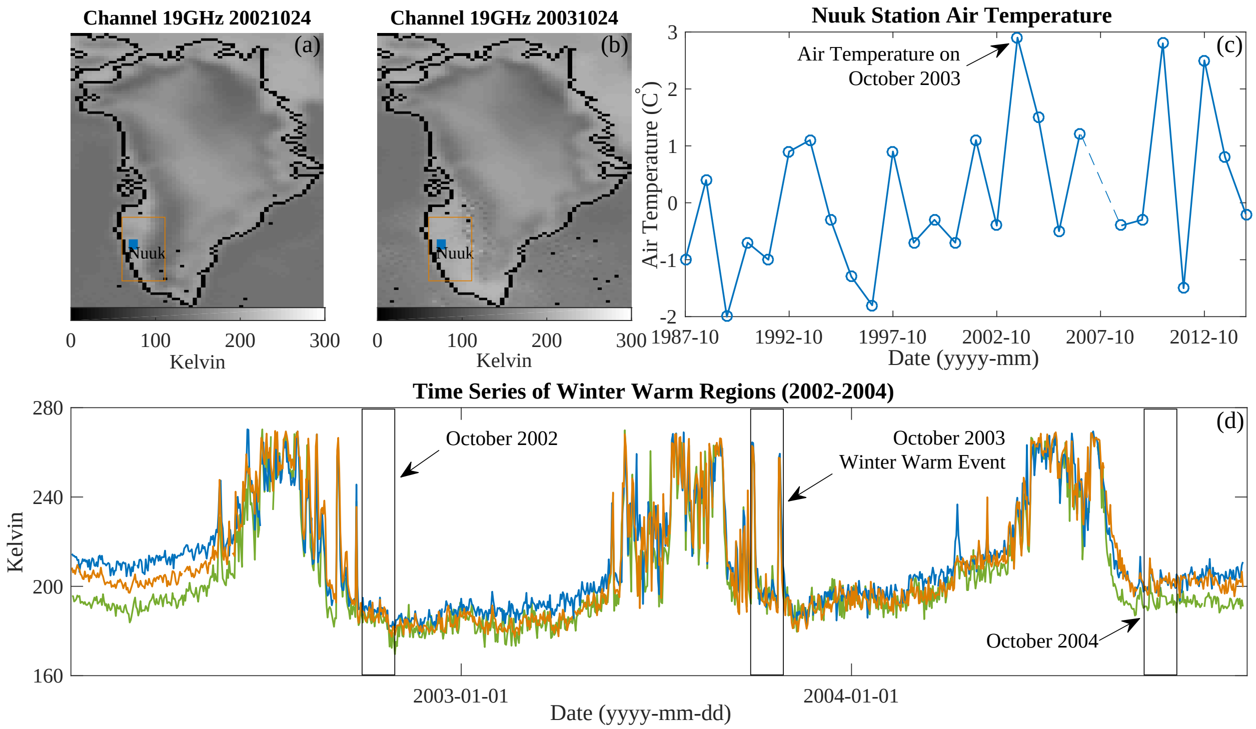

Fig. 19. SSM/I anomalous event: October 2003 winter warm event. (a) and (b) The region inside the box on October 24, 2003 has a higher average brightness temperature than that of the same region on October 24, 2002. (c) The warm event correlated with the air temperature at Nuuk station. (d) The brightness temperature of objects around Nuuk station show a sharp increase during October 2003.

seasonality and noise. The event was found to be related to the Atlantic Subpolar Gyre Warming (southwest to Nuuk at Greenland) in October 2003 (Stein, 2005). Our finding is also consistent with a surface air temperature obtained from the Nuuk station where October 2003 was a record high between 1987 and 2015 (2007 data is missing), as seen in Fig. 19.

Other significant events detected by our framework have been shared with our collaborators for further investigation.

5.3. Computational Efficiency

The computational efficiency of the proposed method was analyzed with two phases: feature extraction and anomaly detection. For the feature extraction process, the computation complexity is $O(tmn)$, where t is the number of images and (m, n) are the number of columns and rows of each image. For anomaly detection, the algorithm complexity is $O(krL)$, where

k is the number of clusters, r is the number of clustering iterations, and L is the length of a time series with extracted features. Table. 2 shows the average processing time for feature extraction on a SSM/I image, and the average processing time for anomaly detection for the longest time series (10,218 days). Since there are missing data in the SSM/I data, the time series have different length, which allows us to evaluate the real algorithm complexity for anomaly detection. Fig. 20(a) shows the anomaly detection time as a function of time series length. The actual computation time is linear and consistent with theoretical complexity. In addition, since there is no dependence among data files during feature extraction and anomaly detection, our method can be easily parallelized using multiple computers or CPUs. For example, Fig. 20(b) shows the different anomaly detection time for a total of 24,356 time series (64 GB) with 1 to 4 CPUs.

Table 2: Computational Efficiency of Feature Extraction and Anomaly Detection

Feature	**Complexity**	**Image Size (m ,n)**	**Process Time (s)**	**I/O (s)**
Extraction	O(tmn)	(304, 448)	0.848	0.171
Anomaly	**Complexity**	**Time Series Length**	**Process Time (s)**	**I/O (s)**
Detection	O(krL)	10218	1.189	0.321

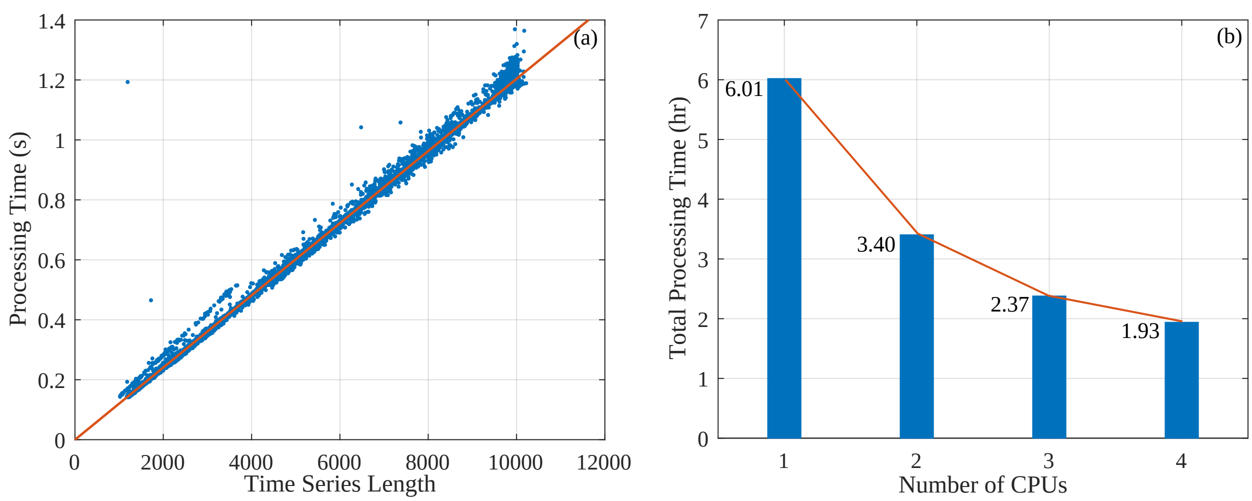

Fig. 20. (a) Illustration of the linear complexity of anomaly detection algorithm. (b) Demonstration of total anomaly detection time with multiple CPUs.

6. Related Work

Anomaly detection has been a topic of active research (Chandola et al., 2009; Gupta et al., 2014; Bhaduri et al., 2011), with the majority of research focusing on point anomaly detection. For example, one-class Kernel Fisher Discriminants (Roth, 2004) was proposed as a method of learning a discriminative boundary close to the normal instances, such that any test instance that does not fall within the learned boundary is considered anomalous. Knorr and Ng (Knox and Ng, 1998; Knorr et al., 2000) developed several distance-based outlier detection algorithms where the core methodology was to score a data instance by counting the number of nearest neighbors that are within a distance d; data instances with the lowest scores were considered outliers. As summarized in Chandola et al. (2009), techniques for detecting point anomalies can be categorized into the following types: classification, nearest neighborhood, clustering, statistical modeling, information, and spectrum based.

Point anomaly detection techniques are extensively used in scientific data, especially in climate research to identify sensor faults and significant events. However, for contextual anomalies that fall within normal data ranges or hide in seasonal patterns, direct use of those classical outlier detection algorithms will fail. Hence, a series of anomaly detection methods leveraging spatial or temporal attributes have been proposed (Chandola et al., 2009; Sun and Chawla, 2004; Alvera-Azcárate et al., 2012). For instance, in the work of Vallis et al., long term time series data was decomposed to remove seasonality before using statistical modeling to find anomalous points (Vallis et al., 2014). Spatial Local Outlier Measure (SLOM) proposed by Sun and Chawla (Sun and Chawla, 2004) can capture the local behavior of datum in its spatial neighborhood. Thus, local spatial outliers can be discovered, which are usually missed by global techniques like "three standard deviations away from the mean". This type of approach handles either temporal or spatial context.

In our work, we detect contextual anomalies in both spatial and temporal contexts. The typical method for spatial-temporal outlier detection consists of three steps (Gupta et al., 2014): (1) Identify spatial objects from the input data. (2) Objects are analyzed to find spatial outliers. (3) Spatial outliers are then verified if they are also temporal outliers. This type of approach sequentially executes spatial and temporal outlier detection, consequently the output is the intersection set of spatial and temporal outliers. For example, Birant and

Kut (Birant and Kut, 2007) proposed a density-based ST-Outlier detection method. They use DBSCAN (Birant and Kut, 2007) to identify spatial outliers first, then validate with their temporal neighbors. If no significant temporal difference was found, the candidate is abandoned. Similarly, Cheng and Li (Cheng and Li, 2006) proposed a four-step method to address the semantic and dynamic properties of geographic phenomena for ST-Outlier detection. However, in order to capture all possible data defects or significant natural events, the union set of spatial and temporal outliers has to be detected. Therefore, we have proposed a single-step ST-Outlier detection algorithm using combined spatial-temporal features, with which we are able to get all the ST-Outliers.

Furthermore, while most of the existing work only detects individual outliers, we aggregate ST-Outliers into anomalous events, which provide more insights into the data as they help reveal underlying processes that may have triggered groups of outliers. This direction of work is related to collective anomalies (Chandola et al., 2009), which have been investigated in some recent work, using statistical models. Das and Neil (Das et al., 2008) used Whats Strange About Recent Events (WSARE) to detect anomalous clusters of counts in categorical data and performed testing to determine if a cluster is a significant anomalous pattern. And a Flexible Genre Model (FGM) was proposed by Xiong et al. (Xiong et al., 2011) to discover anomalous behaviors of groups of points. In contrast to those supervised or semi-supervised methods that assume availability of enough training data with ground truth, our approach is unsupervised and requires no prior knowledge of the data sets.

7. Conclusion

In this work, we have proposed a novel unsupervised contextual anomaly detection framework, which can effectively filter out noisy pixels, discover spatial-temporal outliers, and group those outliers into anomalous events. With this framework, we have successfully identified significant data quality issues and natural events that were subsequently validated by geoscientists. We expect that our experience developing the framework will not only advance anomaly detection in remote sensing but also provide new approaches for speeding up scientific knowledge discovery, especially when combined with interactive data mining and visualization tools. As with any large-scale project, the development is an ongoing effort.

With the support of the National Snow and Ice Data Center (NSIDC), we plan to make our tools publicly available and will continue improving the tools and methodologies, as well as expanding the range of remotely sensed data sets that we can support. We look forward to collaboration and feedback from the community to drive further improvements.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) under grant No.1251257. We thank Garrett Campbell for his advice and assistance during the early stages of this project. We also thank Cathy Fowler and Richard McAllister for their suggestions on user interface design.

References

Alvera-Azcárate, A., et al. Outlier detection in satellite data using spatial coherence. Remote Sensing of Environment 2012;119:84–91.

Andersson, T.. Multivariate statistical analysis. John Wiley and Sons, Inc., New York, 1958.

Bhaduri, K., Matthews, B.L., Giannella, C.R.. Algorithms for speeding up distance-based outlier detection. In: Proceedings of the 17th International Conference on Knowledge Discovery and Data Mining. ACM; 2011. p. 859–867.

Birant, D., Kut, A.. ST-DBSCAN: An algorithm for clustering spatial–temporal data. Data and Knowledge Engineering 2007;60(1):208–221.

Bokhorst, S., Tømmervik, H., Callaghan, T., Phoenix, G., Bjerke, J.. Vegetation recovery following extreme winter warming events in the sub-Arctic estimated using NDVI from remote sensing and handheld passive proximal sensors. Environmental and Experimental Botany 2012;81:18–25.

Borg, E., Fichtelmann, B., Asche, H.. Assessment for remote sensing data: accuracy of interactive data quality interpretation. In: International Conference on Computational Science and Its Applications. Springer; 2011. p. 366–375.

Chandola, V., Banerjee, A., Kumar, V.. Anomaly detection: a survey. ACM Computing Surveys 2009;41(3):15.

Cheng, T., Li, Z.. A multiscale approach for spatio-temporal outlier detection. Transactions in GIS 2006;10(2):253–263.

Chuck, F., Maslanik, J., Haran, T., Scambos, T., Key, J., Emery, W.. AVHRR Polar Pathfinder Twice-daily 5 km EASE-Grid Composites V003, [July 1981 to October 1997]. Boulder, Colorado USA: National Snow and Ice Data Center [January 2016 Accessed] 2000, updated 2007;.

Coumou, D., Rahmstorf, S.. A decade of weather extremes. Nature Climate Change 2012;2(7):491–496.

Das, K., Schneider, J., Neill, D.B.. Anomaly pattern detection in categorical datasets. In: Proceedings of the 14th International Conference on Knowledge Discovery and Data Mining. ACM; 2008. p. 169–176.

Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R., Mearns, L.O.. Climate extremes: observations, modeling, and impacts. Science 2000;289(5487):2068–2074.

Ferguson, C.R., Villarini, G.. Detecting inhomogeneities in the Twentieth Century Reanalysis over the central United States. Journal of Geophysical Research: Atmospheres 2012;117(D5).

Gonzalez, A.R., Datcu, M.. Data cleaning: approach for Earth observation image information mining. In: ESA-EUSC-JRC 2011 Image Information Mining: Geospatial Intelligence from Earth Observation Conference. 2011. p. 117–120.

Gupta, M., Gao, J., Aggarwal, C., Han, J.. Outlier detection for temporal data. Synthesis Lectures on Data Mining and Knowledge Discovery 2014;5(1):1–129.

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D.. Change detection from remotely sensed images: From pixel-based to object-based approaches. ISPRS Journal of Photogrammetry and Remote Sensing 2013;80:91–106.

Isaac, D., Lynnes, C.. Automated data quality assessment in the intelligent archive. White Paper prepared for the Intelligent Data Understanding Program 2003;17.

Knorr, E.M., Ng, R.T., Tucakov, V.. Distance-based outliers: algorithms and applications. The International Journal on Very Large Data Bases 2000;8(3-4):237–253.

Knox, E.M., Ng, R.T.. Algorithms for mining distance based outliers in large datasets. In: Proceedings of the International Conference on Very Large Data Bases. Citeseer; 1998. p. 392–403.

Lee, G., Scott, C.. EM algorithms for multivariate gaussian mixture models with truncated and censored data. Computational Statistics and Data Analysis 2012;56(9):2816–2829.

Li, S., Dragicevic, S., Castro, F.A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., et al. Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing 2016;115:119–133.

López-Moreno, J.I., El-Kenawy, A., Revuelto, J., Azorín-Molina, C., Morán-Tejeda, E., Lorenzo-Lacruz, J., Zabalza, J., Vicente-Serrano, S.M.. Observed trends and future projections for winter warm events in the Ebro basin, northeast Iberian Peninsula. International Journal of Climatology 2014;34(1):49–60.

Ma, Y., Wu, H., Wang, L., Huang, B., Ranjan, R., Zomaya, A., Jie, W.. Remote sensing big data computing: challenges and opportunities. Future Generation Computer Systems 2015;51:47–60.

Maslanik, J., Stroeve., J.. DMSP SSM/I-SSMIS daily polar gridded brightness temperatures, version 4. [July 1987 to June 2015]. Boulder, Colorado USA: National Snow and Ice Data Center Distributed Active Archive Center doi: http://dxdoiorg/105067/AN9AI8EO7PX0 [January 2016 Accessed] 2004, updated 2016;.

Matthes, H., Rinke, A., Dethloff, K.. Recent changes in Arctic temperature extremes: warm and cold spells during winter and summer. Environmental Research Letters 2015;10(11):114020.

McCright, A.M., Dunlap, R.E., Xiao, C.. The impacts of temperature anomalies and political orientation on perceived winter warming. Nature Climate Change 2014;4(12):1077–1081.

Muster, S., Langer, M., Abnizova, A., Young, K.L., Boike, J.. Spatio-temporal sensitivity of MODIS land surface temperature anomalies indicates high potential for large-scale land cover change detection in Arctic permafrost landscapes. Remote Sensing of Environment 2015;168:1–12.

Nghiem, S., Hall, D., Mote, T., Tedesco, M., Albert, M., Keegan, K., Shuman, C., DiGirolamo, N., Neumann, G.. The extreme melt across the Greenland Ice Sheet in 2012. Geophysical Research Letters 2012;39(20).

Rathore, M.M.U., Paul, A., Ahmad, A., Chen, B.W., Huang, B., Ji, W.. Real-time big data analytical architecture for remote sensing application. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2015;8(10):4610–4621.

Roth, V.. Outlier detection with one-class kernel fisher discriminants. In: Advances in Neural Information Processing Systems. 2004. p. 1169–1176.

Rousseeuw, P.J.. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 1987;20:53–65.

Song, X., Wu, M., Jermaine, C., Ranka, S.. Conditional anomaly detection. IEEE Transactions on Knowledge and Data Engineering 2007;19(5):631–645.

Steffen, K., Nghiem, S., Huff, R., Neumann, G.. The melt anomaly of 2002 on the Greenland Ice Sheet from active and passive microwave satellite observations. Geophysical Research Letters 2004;31(20).

Stein, M.. North Atlantic subpolar gyre warming–impacts on Greenland offshore waters. Journal of Northwest Atlantic Fishery Science 2005;36:43–54.

Sun, P., Chawla, S.. On local spatial outliers. In: Fourth International Conference on Data Mining. IEEE; 2004. p. 209–216.

Vallis, O., Hochenbaum, J., Kejariwal, A.. A novel technique for long-term anomaly detection in the cloud. In: 6th USENIX Workshop on Hot Topics in Cloud Computing. 2014. .

Xiong, L., Póczos, B., Schneider, J.G.. Group anomaly detection using flexible genre models. In: Advances in Neural Information Processing Systems. 2011. p. 1071–1079.

List of Figure Captions